for distances between all atoms in the molecule. OMEGA (76) uses a torsion-driven approach for building conformers. It generates low energy conformers for each molecule by assembling it from fragments and searching through possible orientations of the subunit added. WIZARD (77) and COBRA (78), AIMB (79) and MIMUMBA (80) employ artificial intelligence techniques for generating a set of user-specified low energy conformations for a compound. MOLGEO (81) uses a depth-first approach for generating 3D structures based on connectivity using bond length and bond angle tables. IDEALIZE (82) is a molecular mechanics program that minimizes 2D structures to generate the corresponding 3D structure.

2.3.3 2D Pharmacophore Searching. Searching 2D databases is of great importance for accelerating drug discovery. Chemical suppliers provide databases of purchasable compounds that medicinal chemists search for starting material for synthesis or analogs of a lead compound. Different strategies are pursued to search a 2D database to identify compounds of interest. Exact structure search is applied to find out whether a compound is present in the database. Substructure searches identify larger molecules that contain the user-defined query, irrespective of the environment in which the query substructure occurs (83) (Fig. 6.13). Furthermore, substructure searching *can* identify all compounds in a database that share the same core structure. Biochemical data obtained from testing these compounds can be used for generating structure-activity relationships (SARs), even before synthetic plans are made for lead optimization (84). In contrast, superstructure searches are used to find smaller molecules that are embedded in the query (Fig. 6.14). One problem that arises from substructure searches is that the number of compounds identified *can* reach into the thousands. A solution to this problem is ranking the compounds based on similarity to a reference compound. Similarity searches use one or more structural descriptors for quantifying the similarity between compounds in the database and in the query (85, 86) (Fig. 6.15). A review of descriptors used in similarity searches is provided by Willett et al. (86). Beyond structural similarity, activity has also been the subject of several studies. Xue et al. showed that compounds with similar activity could be identified using mini-fingerprints (87–89), physicochemical property descriptors (90), or latent semantic structure indexing (91, 92). In addition, similarity searches *can* be combined with superstructure searches for limiting the number of compounds selected. Flexible match searches are used for identifying compounds that differ from the query structure in user-specified ways. In addition, isomer, tautomer, and parent molecule searches may be done to find in a database isomers, tautomers, or parent molecules of the query.

2.3.4 3D Pharmacophores

2.3.4.1 Ligand-Based Pharmacophore Generation. Ligand-based pharmacophores are typically used when the crystallographic, solution structure, or modeled structure of a protein cannot be obtained. When a set of active compounds is known and it is hypothesized that all compounds bind in a similar way to the protein, then common groups should interact with the same protein residues. Thus, a pharmacophore capturing these common features should be able to identify from a database novel compounds that bind to the same site of the protein as the known compounds do. The process of deriving a pharmacophore, called ***pharmacophore mapping,*** consists of three steps: (1) identifying common binding elements that are responsible for biological activity; (2) generating potential conformations that active compounds may adopt; and (3) determining the 3D relationship between pharmacophore elements in each conformation generated. To build a pharmacophore based on a set of active compounds, two methods are usually applied. One method is to generate a set of minimum energy conformations for each ligand and search for common structural features. Another method is to consider all possible conformations of each ligand to evaluate shared orientations of common functional groups. Analyzing many low energy conformers of active compounds can suggest a range of the distance between key groups that will take in account the flexibility of the ligands and of the protein. This task can be performed either manually or automatically.